

Joint DOA and Frequency Estimation with Sub-Nyquist Sampling

Liang Liu and Ping Wei

Abstract—In this paper, to jointly estimate the frequency and the direction-of-arrival(DOA) of the narrowband far-field signals, a novel array receiver architecture is presented by the concept of the sub-Nyquist sampling techniques. In particular, our contribution is threefold. i) First, we propose a time-space union signal reception model for receiving array signals, where the sub-Nyquist sampling techniques and arbitrary array geometries are employed to decrease the time-domain sampling rate and improve the DOA estimation accuracy. A better joint estimation is obtained in the higher time-space union space. ii) Second, two joint estimation algorithms are proposed for the receiving model. One is based on a trilinear decomposition from the third-order tensor theory and the other is based on subspace decomposition. iii) Third, we derive the corresponding Cramér-Rao Bound (CRB) for frequency and DOA estimates. In the case of the branch number of our architecture is equal to the reduction factor of the sampling rate, it is observed that the CRB is robust in terms of the number of signals, while the CRB based on the Nyquist sampling scheme will increase with respect to the number of signals. In addition, the new steer vectors of the union time-space model are completely uncorrelated under the limited number of sensors, which improves the estimation performance. Furthermore, the simulation results demonstrate that our estimates via the receiver architecture associated with the proposed algorithms closely match the CRB according to the noise levels, the branch number and the source number as well.

Index Terms—Direction-of-arrival estimation, frequency estimation, sub-Nyquist sampling, Cramér-Rao Bound.

I. INTRODUCTION

JOINT estimation of carrier frequency and direction of arrival (DOA) for multiple signals is desired in many practical applications. For example, Cognitive Radio (CR) technique might be a good way to cope with the problem of the spectral congestion [1]–[5]. One of the most important functions of CRs is to detect locally idle spectrum and then make the spectrum access from the concept of spectrum sensing. Generally, there are three dimensions of spectrum space, i.e., time, frequency and space. With the development of array processing techniques [6]–[8], the spatial spectrum or DOA of a signal can be thought as a new approach to improve the performance of CRs. Therefore, more effort have been spent on how to jointly estimate carrier frequencies and their DOAs of multiple signals [9], [10]. Unfortunately, both of them exist at least two shortcomings. One is the pair matching problem for the carrier frequencies associated with the DOAs. The other is that time domain sampling rate is

equal to or larger than the Nyquist sampling rate which is considered as a bottleneck for wideband signal processing by CRs. For instance, it leads to prohibitive Nyquist sampling rate and massive sampling data to be processed if the spectrum needed to be monitored from 300 MHz to several GHz [1]–[5].

To deal with the problem of high sampling rate, recently, the sub-Nyquist sampling technique has been proposed to reconstruct a multiband signal from the data obtained under the Nyquist sampling rate [11]–[14]. Inspired by the idea, some methods were presented for the joint estimation of DOA and carrier frequency based on sub-Nyquist sampling rates. The authors of [15] suggested a new structure, where each output of a linear array is carried out through the multi-coset sampling. In [15] the minimum redundancy array (MRA) is employed to estimate the DOA of more uncorrelated sources than sensors. In this way, the wide-sense stationary signal can be compressed in both the time domain and the spatial domain. The frequency and DOA estimation accuracy are limited by the reciprocal of block length and array aperture, respectively. And it need a two-dimensional (2D) peak searching to get the frequency and DOA estimation from the 2D power spectrum. To simplify the hardware complexity, an additional identical delayed channel for each antenna is suggested in [16]. Herein, the problem of pairing ambiguity will arise using an underlying uniform linear array (ULA). And then [17] proposed a structure, which has the same hardware complexity as that of [16]. In [18], the authors proposed the so-called space-time array to jointly estimate frequency and DOA when the number of sources is more than the number of sensors. However, those methods in [17], [18] are limited to ULA because they make use of the rotation invariance property of ULA. More recently, two joint DOA and carrier frequency recovery approaches based on the L-shaped ULAs are presented in [19]. However, all of these papers did not give a unified signal reception model for the array receivers.

Dealing with the problem of joint frequency and DOA estimation, it is widespread that the spatial samplings are less because of limited sensors number, and the temporal samplings are enough. A kind of very natural viewpoint is jointly considering in both time domain and space domain. If we unite time and space domain through elaborately modeling, we will have more chances to classify targets. Because the differences between vectors from not only spatial space but also temporal space will be reserved in the new union space, besides, the differences between vectors in the new union space will be enlarged even if the differences between vectors from any one space are small, we employ Kronecker or Khatri-Rao product to jointly estimate the frequency and DOA from

The authors are with the Center for Cyber Security, School of Electronic Engineering, University of Electronic Science and Technology of China, Chengdu, 611731 China (e-mail:liu_yinliang@outlook.com; pwei@uestc.edu.cn).

the time-space union space in this paper. Because the new vectors in the union space have bigger differences and less correlation, we can better classify the targets based on the differences from not only frequency domain but also spatial domain.

In this paper, a new array receiver architecture is proposed. Associated with two sub-Nyquist sampling based methods, we simultaneously estimate the frequencies and DOAs of multiple narrowband far-field signals impinging on a array, where signals' carrier frequencies spread around the whole wide spectrum. It is noteworthy that the array is not limited to ULA. The other arrays can be applied to gain their advantage, such as MAR can achieve a higher estimation accuracy than ULA with same sensor number. Since the reception model of our receiver makes use of the result on Kronecker product, the joint DOA and frequency estimation will benefit from it. In addition, the Cramér-Rao Bound (CRB) for spatial phase estimation is also derived based on this model. It is proven that the CRB is not affected by the signal number when the branch number of our architecture is equal to the sampling rate reduction factor, while the CRB using Nyquist sampling will increase according to the signal number. In other words, our model's CRB is lower than the CRB which employs Nyquist sampling. Finally, the simulations confirm the above conclusion on CRB and the superior performance of the proposed methods from three aspects: noise level, the number of branch, and the number of source as well.

This paper is organized as follows: in Section II, we describe the basic array signal model and point out the objective of this paper. In Section III, the proposed receiver architecture is presented, and a new signal reception model is derived. In Section IV, two joint DOA and frequency estimation methods for the receiver architecture are proposed. In Section V, we deduce the corresponding CRB and demonstrate the result on CRB. Section VI carries out the simulation experiment and finally the conclusions of this paper are given in Section VII.

The following notations are used in the paper. $(\cdot)^T$, $(\cdot)^H$, and $(\cdot)^\dagger$ denote the transpose, Hermitian transpose, and Moore-Penrose pseudo-inverse, respectively. $E(\cdot)$ stands for the expectation operator. x_j is the j th entry of a vector \mathbf{x} . \mathbf{A}_i , \mathbf{A}^j , and A_{ij} are the i th row, the j th column, and (i, j) th entry of a matrix \mathbf{A} , respectively. \otimes , \odot , and $*$ denote the Kronecker product, Hadamard product, and Khatri-Rao product, respectively. \mathbf{I}_M stands for an $M \times M$ identity matrix.

II. SIGNAL MODEL AND OBJECTIVE

In this section, we will give the array signal model and fundamental assumptions as well as the objective of this paper.

A. Array signal model

Consider K narrowband far-field signals impinging on an array composed of M ($M > K$) sensors. It is assumed that the signals' center frequencies are separate widely. Thus, the narrowband far-field signals can be modeled as multiband signals in [14]. The array output can be written as [6]

$$\mathbf{x}(t) = \mathbf{A}\mathbf{s}(t) + \mathbf{n}(t), \quad (1)$$

where $\mathbf{x}(t) = [x_1(t), \dots, x_M(t)]^T$ is the measurement vector, $\mathbf{s}(t) = [s_1(t), \dots, s_K(t)]^T$ is the vector of all signal values, where the signals are uncorrelated, $\mathbf{n}(t) = [n_1(t), \dots, n_M(t)]^T$ is the zero-mean complex spatially and temporarily white Gaussian noise vector, whose variance is σ^2 . As the most widely used array, who also derive many non-uniform linear array, the ULA is taken into consideration. The array manifold matrix has the form as $\mathbf{A} = [\mathbf{a}(\phi_1), \dots, \mathbf{a}(\phi_K)]$, where array steering vector $\mathbf{a}(\phi_k) = [\exp(j\phi_k 0), \dots, \exp(j\phi_k (M-1))]^T$, and spatial phase

$$\phi_k = \frac{\pi d \sin(\theta_k) f_k}{f_N}, \quad (2)$$

where d is the distance between two consecutive antennas in half-wavelengths corresponding to the Nyquist sampling rate f_N , θ_k and f_k are the DOA and the center frequency of $s_k(t)$, respectively. The sensor position vector is $\mathbf{d} = [0, 1, \dots, M-1]d$. Note that the array is not limited to ULA for our receiver architecture and algorithms in the next few sections.

The frequency domain output can be written as

$$\mathbf{X}(f) = \mathbf{A}\mathbf{S}(f) + \mathbf{N}(f), \quad (3)$$

where $\mathbf{X}(f) = [X_1(f), \dots, X_M(f)]^T$, $\mathbf{S}(f) = [S_1(f), \dots, S_K(f)]^T$, and $\mathbf{N}(f) = [N_1(f), \dots, N_M(f)]^T$ are the frequency domain expression of $\mathbf{x}(t)$, $\mathbf{s}(t)$, $\mathbf{n}(t)$, respectively. $X_m(f)$ is the Fourier transform of $x_m(t)$.

B. Objective statement

The objective of this paper is to simultaneously estimate the carrier frequency f_k and DOA θ_k of multiple signals $s_k(t)$. To achieve this goal, we will introduce the novel methods under the Nyquist sampling rate as follows.

III. PROPOSED RECEIVER ARCHITECTURE AND SIGNAL RECEPTION MODEL

Now, we modify the traditional array signal receiver architecture and introduce the sub-Nyquist sampling technique into the architecture to reduce sampling rate. In this section, a novel architecture is presented and the corresponding signal reception model is derived.

A. Receiver architecture

Our receiver architecture is shown in Fig.1. We apply the multi-coset sampling [14] in Fig.1 as representative of sub-Nyquist sampling technology. In Fig. 1, there are M sensors and every sensor is followed by P delay branches. All the ADCs are well-synchronized and sample at a sub-Nyquist sampling rate of $f_{sub} = f_N/L$, where $f_N = 1/T_N$ is the Nyquist sampling rate and L is the sampling rate reduction factor, where T_N is Nyquist sampling interval. The constant set $C = [c_1, \dots, c_P]$ is referred to the sampling pattern where $0 \leq c_1 < c_2 < \dots < c_P \leq L-1$. $y_{mp}[n]$ denotes the sampled signal corresponding to the m th sensor, p th branch.

The average sampling rate of the multi-coset sampling is

$$f_E = \frac{P f_N}{L}, \quad (4)$$

which is lower than the Nyquist rate f_N when $P < L$.

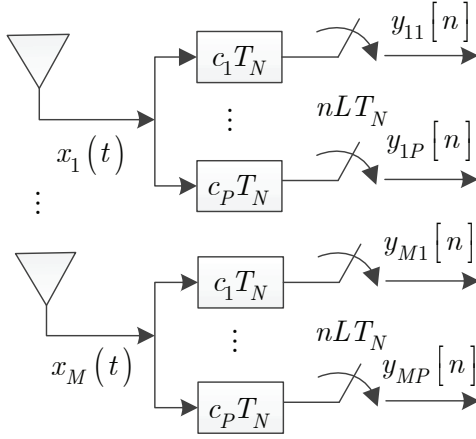


Fig. 1. Proposed receiver architecture.

B. Signal reception model

According to the conclusion of [14], the relationship between the discrete-time Fourier transform $Y_{mp}(e^{j2\pi f T_N})$ of the signal $y_{mp}[n]$ and the Fourier transform $X_m(f)$ of $x_m(t)$ is as follows.

$$Y_{mp}(e^{j2\pi f T_N}) = \frac{1}{LT_N} \sum_{l=1}^L \exp\left(j\frac{2\pi}{L} c_p l\right) X_{ml}(f),$$

$$0 \leq p \leq P, f \in \mathcal{F} \triangleq [0, f_{sub}). \quad (5)$$

The matrix form of (5) is expressed as

$$\mathbf{Y}_m(f) = \mathbf{B} \bar{\mathbf{X}}_m(f), f \in \mathcal{F}, 1 \leq m \leq M, \quad (6)$$

where $\mathbf{Y}_m(f) = [Y_{m1}(e^{j2\pi f T_N}), \dots, Y_{mP}(e^{j2\pi f T_N})]^T$, $\bar{\mathbf{X}}_m(f) = [X_{m1}(f), \dots, X_{mL}(f)]^T$, $X_{ml}(f) = X_m(f + (l-1)f_{sub})$, $B_{il} = \frac{1}{LT_N} \exp(j\frac{2\pi}{L} c_i l)$. For convenience, we multiply both sides of (6) by $\sqrt{LT_N}$ to normalize the row vectors of \mathbf{B} . Then redefine $B_{il} \triangleq \frac{1}{\sqrt{L}} \exp(j\frac{2\pi}{L} c_i l)$, $Y_{mp}(f) \triangleq \sqrt{LT_N} Y_{mp}(e^{j2\pi f T_N})$. From (3),

$$X_{ml}(f) = \mathbf{A}_m \hat{\mathbf{S}}_l(f) + N_{ml}(f), 1 \leq l \leq L, f \in \mathcal{F}, \quad (7)$$

where $\hat{\mathbf{S}}_l(f) = [S_{1l}(f), \dots, S_{KL}(f)]^T$, $S_{kl}(f) = S_k(f + (l-1)f_{sub})$, $N_{kl}(f) = N_k(f + (l-1)f_{sub})$. In matrix form, (7) can be expressed as

$$\bar{\mathbf{X}}_m(f) = (\mathbf{I}_L \otimes \mathbf{A}_m) \hat{\mathbf{S}}(f) + \hat{\mathbf{N}}_m(f), f \in \mathcal{F}, \quad (8)$$

where $\hat{\mathbf{S}}(f) = [\hat{\mathbf{S}}_1^T(f), \dots, \hat{\mathbf{S}}_L^T(f)]^T$, $\hat{\mathbf{N}}_m(f) = [N_{m1}(f), \dots, N_{mL}(f)]^T$.

Substituting (8) into (6), we get

$$\begin{aligned} \mathbf{Y}_m(f) &= \mathbf{B} (\mathbf{I}_L \otimes \mathbf{A}_m) \hat{\mathbf{S}}(f) + \mathbf{B} \hat{\mathbf{N}}_m(f) \\ &= (\mathbf{A}_m \otimes \mathbf{B}) \bar{\mathbf{S}}(f) + \mathbf{B} \hat{\mathbf{N}}_m(f), f \in \mathcal{F}, 1 \leq m \leq M, \end{aligned} \quad (9)$$

where $\bar{\mathbf{S}}(f) = [\bar{\mathbf{S}}_1^T(f), \dots, \bar{\mathbf{S}}_K^T(f)]^T$, $\bar{\mathbf{S}}_k(f) = [S_{k1}(f), \dots, S_{kL}(f)]^T$.

Then, combining all m can result in

$$\mathbf{Y}(f) = (\mathbf{A} \otimes \mathbf{B}) \bar{\mathbf{S}}(f) + (\mathbf{I}_M \otimes \mathbf{B}) \hat{\mathbf{N}}(f) \quad (10)$$

$$\triangleq \mathbf{G} \bar{\mathbf{S}}(f) + \mathbf{I}_B \hat{\mathbf{N}}(f), f \in \mathcal{F}, \quad (11)$$

where $\mathbf{Y}(f) = [\mathbf{Y}_1^T(f), \dots, \mathbf{Y}_M^T(f)]^T$, $\hat{\mathbf{N}}(f) = [\hat{\mathbf{N}}_1^T(f), \dots, \hat{\mathbf{N}}_M^T(f)]^T$. Actually, $\mathbf{Y}(f)$ in (10) is the matrix form of the output of all branches of all sensors.

Because $s_k(t)$ is a narrowband signal, there is only one frequency band which is occupied in $\bar{\mathbf{S}}_k(f)$. Further, $\bar{\mathbf{S}}_k(f)$ is a sparse vector of length L when k is fixed and there is only one index (marked as l_k), which is activated. Since it is assumed that those signals' carrier frequencies are far between, any two signals are not in the same sub-band. Namely, l_i is not equal to l_j for any $i \neq j$. $\Omega = [l_1, \dots, l_K]$ denotes the activated index set of $\bar{\mathbf{S}}_k(f)$ and \mathbf{B} . Further, $\bar{\mathbf{S}}(f)$ is a K -sparse vector of length KL . The support index \mathcal{S} of $\bar{\mathbf{S}}(f)$ and \mathbf{G} is determined as

$$\mathcal{S}_k = (k-1)L + l_k. \quad (12)$$

With the knowledge of \mathcal{S} , (11) can be written as

$$\mathbf{Y}(f) = \mathbf{G}_{\mathcal{S}} \bar{\mathbf{S}}^{\mathcal{S}}(f) + \mathbf{I}_B \hat{\mathbf{N}}(f) \quad (13)$$

$$= (\mathbf{A} * \mathbf{B}_{\Omega}) \bar{\mathbf{S}}^{\mathcal{S}}(f) + \mathbf{I}_B \hat{\mathbf{N}}(f), f \in \mathcal{F}. \quad (14)$$

Remark 1. It is clear that (6) and (7) are the sub-Nyquist sampling model and DOA model, respectively. If we just severally consider the frequency estimation and DOA estimation in (6) and (7), we will meet the match problem and can not comprehensively classify the targets. The target classification can be jointly handled in a union space based on equation (13), where the compressive sampling can be applied in time and space domain, respectively. As shown in Fig.2 and explained in section I, we have more chances and better performance to classify the targets in the time-space union domain. In terms of the performance, we will further study in Section V. On the contrary, [15] gives the 2D power spectrum instead of reception model. In [16], [17], the sub-Nyquist sampling is also applied to array receiver, but only separate models are given. Besides, those methods in [17], [18] are limited to ULA because they make use of the rotation invariance property of ULA. Their receiver can not employ other particular array to make use of those arrays' advantages.

In (10), since the Khatri-Rao product is used to unify the frequency domain and spatial domain into a two-dimensional matrix form, this equation can be viewed from the perspective of third-order tensor. We will discuss the third-order tensor in the next section. On the other hand, if \mathbf{G} is regarded as a special array manifold, the subspace decomposition theory can be employed. Of course, the different perspectives will derive different methods, which will be analyzed in detail.

IV. JOINT DOA AND FREQUENCY ESTIMATION ALGORITHM

A. Algorithm based on trilinear decomposition

It is easy to express (13) in element form as

$$Y_{mp}(f) = \sum_{k=1}^K A_{mk} B_{plk} S_{klk}(f), f \in \mathcal{F}. \quad (15)$$

From (15), $Y_{mp}(f)$ can be regarded as a third-order tensor and the problem can be solved by trilinear decomposition [20], [21]. This problem is different from the standard trilinear decomposition problem since \mathbf{B} is known here. Even so, we do not know which columns are activated. So, we can use the standard trilinear decomposition algorithm, such as alternating least squares (ALS) [22] and regularized alternating least squares (RALS) [23], [24], where some sufficient conditions for uniqueness up to permutation and scalings of the decomposition are provided. After the decomposition, we can obtain $\tilde{\mathbf{A}}$, $\tilde{\mathbf{B}}$, and $\tilde{\mathbf{S}}(f)$.

Since not only every column vector of $\tilde{\mathbf{A}}$ but also every row vector of $\tilde{\mathbf{S}}(t)$ (the inverse discrete-time Fourier transform of $\tilde{\mathbf{S}}(f)$) can be viewed as a single tone, periodogram is applied on every column vector of $\tilde{\mathbf{A}}$ or every row vector of $\tilde{\mathbf{S}}(t)$ to achieve ϕ or f maximum likelihood (ML) estimation [25]. The periodogram is briefly introduced in the following. For an N -length single frequency sine wave $z(n) = \exp(j\omega_0 n)$, $n = 1, 2, \dots, N$, the ML estimation for ω_0 is realized through

$$\hat{\omega}_0 = \arg \max_{\omega \in [0, 2\pi)} \left| \sum_{n=1}^N z(n) \exp(-j\omega n) \right|. \quad (16)$$

Similarly, for an arbitrary array form $\mathbf{z} = \mathbf{a}(\phi_0)$, the ML estimation for ϕ_0 is realized by

$$\hat{\phi}_0 = \arg \max_{\phi \in [0, 2\pi)} |\mathbf{a}^H(\phi) \mathbf{z}|. \quad (17)$$

We determine Ω through comparing the correlation coefficient of the column between \mathbf{B} and $\tilde{\mathbf{B}}$ as

$$\Omega_k = \arg \max_j r_{kj} = \left| \frac{(\tilde{\mathbf{B}}^k)^H \mathbf{B}^j}{\|\tilde{\mathbf{B}}^k\| \|\mathbf{B}^j\|} \right|, j = 1, \dots, L. \quad (18)$$

The received signal's frequency estimation \bar{f}_k is obtained by applying periodogram to $\tilde{\mathbf{S}}_k(t)$. Besides, there is a relationship between \bar{f}_k and the original signal's frequency f_k :

$$f_k = (\Omega_k - 1) \frac{f_N}{L} + \bar{f}_k. \quad (19)$$

And then, once ϕ_k and f_k are known, θ_k can be acquired by (2). We outline the main steps of this method named joint algorithm based on trilinear decomposition (JDFTD) in table I.

TABLE I
ALGORITHM JDFTD

1)	Obtain $\tilde{\mathbf{A}}$, $\tilde{\mathbf{B}}$, and $\tilde{\mathbf{S}}(f)$ using RALS according to (15);
2)	Gain ϕ_k applying (16) or (17) to \mathbf{A} ;
3)	Determine Ω_k according to (18);
4)	Get $\tilde{\mathbf{S}}_k(t)$ and \bar{f}_k by applying IFFT, (16) to $\tilde{\mathbf{S}}_k(f)$, $\tilde{\mathbf{S}}_k(t)$, consecutively;
5)	Compute f_k through (19);
6)	Calculate θ_k through (2);

B. Algorithm based on subspace decomposition

In this subsection, we will take advantage of subspace decomposition theory [7]. The covariance matrix of $\mathbf{Y}(f)$, $f \in \mathcal{F}$ is given by

$$\mathbf{R} = \mathbf{E}(\mathbf{Y}(f) \mathbf{Y}^H(f)) = \mathbf{G}_S \mathbf{R}_{\tilde{\mathbf{S}}} \mathbf{G}_S^H + \sigma^2 \mathbf{I}_{MP}, \quad (20)$$

where $\mathbf{R}_{\tilde{\mathbf{S}}}$, $\sigma^2 \mathbf{I}_{MP}$ are the source and noise covariance matrix, respectively. (20) makes use of $\mathbf{I}_{MP} = \mathbf{I}_B \mathbf{I}_B^H$. In actual situation, we can obtain the estimate of the autocovariance matrix through

$$\mathbf{R} = \frac{1}{T/L} \sum_{f=1}^{T/L} \mathbf{Y}(f) \mathbf{Y}^H(f) \quad (21)$$

when T the snapshots of observation are sufficient.

Applying the singular value decomposition (SVD) to \mathbf{R} results in

$$\mathbf{R} = \mathbf{U}_S \mathbf{D}_S \mathbf{U}_S^H + \mathbf{U}_N \mathbf{D}_N \mathbf{U}_N^H, \quad (22)$$

where \mathbf{U}_S and \mathbf{U}_N are signal subspace and noise subspace, respectively. Since the signal subspace and the noise subspace are orthogonal,

$$\mathbf{a}_l(\phi) \perp \mathbf{U}_N \quad (23)$$

holds, where $\mathbf{a}_l(\phi) = \mathbf{a}(\phi) * \mathbf{B}_l$. Computing the pseudo-spectra

$$P(l, \phi) = \frac{1}{\|\mathbf{a}_l^H(\phi) \mathbf{U}_N\|^2} \quad (24)$$

and applying a peak search algorithm, ϕ_k , l_k are obtained. Further, we have \mathbf{A} , Ω . Since \mathcal{S} can be solved by (12), the least square solution of $\tilde{\mathbf{S}}^{\mathcal{S}}(f)$ is given by

$$\tilde{\mathbf{S}}^{\mathcal{S}}(f) = \mathbf{G}_S^\dagger \mathbf{Y}(f), f \in \mathcal{F}, \quad (25)$$

where $\mathbf{G}_S = \mathbf{A} * \mathbf{B}_\Omega$. Similarly, f_k and θ_k can be calculated through the step 4)-6) of Table I. We outline the main steps of this method named joint algorithm based on subspace decomposition (JDFSD) in table II.

V. CRAMÉR-RAO BOUND

In this section, we derive the CRB on the covariance matrix of any unbiased estimator of ϕ . If the signal autocorrelation matrices are defined as

$$\mathbf{R}_S \triangleq \frac{1}{T} \sum_{f=1}^T \mathbf{S}(f) \mathbf{S}(f)^H \quad (26)$$

$$\mathbf{R}_{\tilde{\mathbf{S}}} \triangleq \frac{1}{T/L} \sum_{f=1}^{T/L} \tilde{\mathbf{S}}^{\mathcal{S}}(f) \tilde{\mathbf{S}}^{\mathcal{S}H}(f) \quad (27)$$

TABLE II
ALGORITHM JDFSD

1)	Calculate \mathbf{R} according to (21);
2)	Gain \mathbf{U}_N by applying SVD to \mathbf{R} ;
3)	Compute $P(l, \phi)$ through (24);
4)	Acquire ϕ_k, l_k by peak search algorithm, further, we have \mathbf{A} , Ω , and \mathbf{G}_S ;
5)	Determine $\bar{\mathbf{S}}^S(f)$ according to (25);
6)	Execute the step 4)-6) of Table I.

and according to the exchangeability of summing and the definition of $\bar{\mathbf{S}}^S(f)$,

$$\mathbf{R}_S = \frac{1}{L} \mathbf{R}_{\bar{S}} \quad (28)$$

holds.

Based on the form of \mathbf{B} , $\mathbf{R}_B = \mathbf{I}_B \mathbf{I}_B^H = \mathbf{I}_{MP}$ holds. According to model (10), the log-likelihood function of the data $\mathbf{Y}(f)$, $f \in \mathcal{F}$ is given by

$$\ln L = \text{const} - \frac{1}{\sigma^2} \sum_{f=1}^{T/L} \left(\mathbf{Y}(f) - \mathbf{G}_S \bar{\mathbf{S}}^S(f) \right)^H \cdot \left(\mathbf{Y}(f) - \mathbf{G}_S \bar{\mathbf{S}}^S(f) \right). \quad (29)$$

Comparing (29) with APPENDIX E.1 in [26], and making use of the conclusion of Section IV equation (4.6) in [26], the CRB of our model is given by

$$\begin{aligned} \text{CRB}_{\text{sub}} &= \frac{\sigma^2}{2T/L} \left(\Re \left((\mathbf{E}^H \mathbf{P}_{\mathbf{G}_S} \mathbf{E}) \odot \mathbf{R}_{\bar{\mathbf{S}}}^H \right) \right)^{-1} \\ &= \frac{\sigma^2}{2T} \left(\Re \left((\mathbf{E}^H \mathbf{P}_{\mathbf{G}_S} \mathbf{E}) \odot \mathbf{R}_{\bar{\mathbf{S}}}^H \right) \right)^{-1} \end{aligned} \quad (30)$$

where $\mathbf{P}_{\mathbf{G}_S} = \mathbf{I} - \mathbf{G}_S \mathbf{G}_S^\dagger$, where $\mathbf{G}_S^\dagger = (\mathbf{G}_S^H \mathbf{G}_S)^{-1} \mathbf{G}_S^H$, $\mathbf{E} = [\mathbf{E}_1, \dots, \mathbf{E}_K]$, $\mathbf{E}_i = \frac{d\mathbf{G}_{S_i}}{d\phi_i}$. For comparing, the CRB which employs the Nyquist sampling (marked as CRB_{Ny}) [26] is rewritten here.

$$\text{CRB}_{Ny} = \frac{\sigma^2}{2T} \left(\Re \left((\mathbf{D}^H \mathbf{P}_A \mathbf{D}) \odot \mathbf{R}_S^H \right) \right)^{-1} \quad (31)$$

where $\mathbf{P}_A = \mathbf{I} - \mathbf{A} \mathbf{A}^\dagger$, $\mathbf{D} = [\mathbf{D}_1, \dots, \mathbf{D}_K]$, $\mathbf{D}_i = \frac{d\mathbf{A}_i}{d\phi_i}$.

Next, we will show that CRB_{sub} is lower than CRB_{Ny} when the branch number of our architecture is equal to the sampling rate reduction factor ($P = L$). At this point, on the one hand, the sub-Nyquist sampling and Nyquist sampling obtain equal snapshot in the same time, on the other hand, the received data by sub-Nyquist sampling can just be viewed as the rearrangement of the received data by Nyquist sampling. The proof will carry out in two steps: first, CRB_{Ny} will increase with the number of source K ; secondly, CRB_{sub} is not influenced by the number of source, and the CRB_{sub} is equal to the minimum value of CRB_{Ny} .

For convenience, let us introduce the following notation: $\mathbf{A}_+ = [\mathbf{A}, a]$, $\mathbf{D}_+ = [\mathbf{D}, d]$, $\mathbf{T}_+ = \mathbf{D}_+^H (\mathbf{I} - \mathbf{A}_+ \mathbf{A}_+^\dagger) \mathbf{D}_+$, $\mathbf{T} = \mathbf{D}^H (\mathbf{I} - \mathbf{A} \mathbf{A}^\dagger) \mathbf{D}$, where a is the steer vector corresponding to the increased angle ϕ_{K+1} , $d = \frac{da}{d\phi_{K+1}}$. $\mathbf{R}_+ = \begin{bmatrix} \mathbf{R}_S & \mu \\ \mu^H & \nu \end{bmatrix}$ is the covariance matrix of all $K+1$ signals,

where \mathbf{R}_S is the covariance matrix of original K signals, μ is cross-correlation vector between the increased signal and the original signals, ν is average power of the increased signal. $(\mathbf{R})_K$ denotes the K th order principal minor determinant of \mathbf{R} .

Making use of the nested structure of \mathbf{A}_+ and the matrix inversion lemma [27],

$$\mathbf{A}_+ \mathbf{A}_+^\dagger = \mathbf{A} \mathbf{A}^\dagger + \mathbf{U}, \quad (32)$$

holds after some matrix manipulations, where $\mathbf{U} = \frac{1}{a^H (\mathbf{I} - \mathbf{A} \mathbf{A}^\dagger) a} ((\mathbf{I} - \mathbf{A} \mathbf{A}^\dagger) a) ((\mathbf{I} - \mathbf{A} \mathbf{A}^\dagger) a)^H \succeq 0$ since $(\mathbf{I} - \mathbf{A} \mathbf{A}^\dagger)$ is a projection matrix.

Taking (32) and the nested structure of \mathbf{D}_+ into \mathbf{T}_+ leads to

$$(\mathbf{T}_+)_K = \mathbf{T} - \mathbf{D}^H \mathbf{U} \mathbf{D}. \quad (33)$$

It is easy to proof that $\mathbf{D}^H \mathbf{U} \mathbf{D} \succeq 0$ with $\mathbf{U} \succeq 0$. We thus hold

$$(\mathbf{T}_+)_K = \mathbf{T} - \mathbf{D}^H \mathbf{U} \mathbf{D} \preceq \mathbf{T}. \quad (34)$$

To proceed, we give a proposition.

Proposition 1. For a Hermitian matrix $\mathbf{M}_+ \in C^{(K+1)(K+1)}$, if $\mathbf{M}_+ \succeq 0$, then $(\mathbf{M}_+^{-1})_K \succeq ((\mathbf{M}_+)_K)^{-1}$.

Proof: Since $\mathbf{M}_+ \succeq 0$ it can be decomposed as $\mathbf{M}_+ = \mathbf{H}_+ \mathbf{H}_+^H$, where $\mathbf{H}_+ \in C^{K \times K}$ [27]. Dividing \mathbf{H}_+ into blocks as $\mathbf{H}_+ = [\mathbf{H}, h]$ results in

$$\mathbf{M}_+ = \begin{bmatrix} \mathbf{H}^H \mathbf{H} & \mathbf{H}^H h \\ h^H \mathbf{H} & h^H h \end{bmatrix} = \begin{bmatrix} \mathbf{M} & \mathbf{H}^H h \\ \mathbf{H}^H h & h^H h \end{bmatrix}. \quad (35)$$

Making use of the matrix inversion lemma [27],

$$(\mathbf{M}_+^{-1})_K = \mathbf{M}^{-1} + \mathbf{V} \succeq \mathbf{M}^{-1} = ((\mathbf{M}_+)_K)^{-1} \quad (36)$$

holds since $\mathbf{V} = \frac{1}{h^H (\mathbf{I} - \mathbf{H} \mathbf{H}^H) h} (\mathbf{M}^{-1} \mathbf{H}^H h) (\mathbf{M}^{-1} \mathbf{H}^H h)^H \succeq 0$. ■

Note that $\mathbf{T}_+ = \mathbf{D}_+^H (\mathbf{I} - \mathbf{A}_+ \mathbf{A}_+^\dagger) \mathbf{D}_+ \succeq 0$, $\mathbf{R}_+ \succeq 0$, $\Re(\mathbf{T}_+ \odot \mathbf{R}_+^H) \succeq 0$ holds. Making use of proposition 1, we then obtain

$$\begin{aligned} \left(\Re(\mathbf{T}_+ \odot \mathbf{R}_+^H) \right)^{-1}_K &\succeq \left(\Re(\mathbf{T}_+ \odot \mathbf{R}_+^H) \right)^{-1} \\ &= \left(\Re((\mathbf{T}_+)_K \odot \mathbf{R}_S^H) \right)^{-1} \\ &\succeq \left(\Re(\mathbf{T} \odot \mathbf{R}_S^H) \right)^{-1}. \end{aligned} \quad (37)$$

Considering $\text{CRB}_{Ny+} = \frac{\sigma^2}{2T} \left(\Re(\mathbf{T}_+ \odot \mathbf{R}_+^H) \right)^{-1}$, $\text{CRB}_{Ny} = \frac{\sigma^2}{2T} \left(\Re(\mathbf{T} \odot \mathbf{R}_S^H) \right)^{-1}$, we get

$$\begin{aligned} (\text{CRB}_{Ny+})_K &= \frac{\sigma^2}{2T} \left(\left(\Re(\mathbf{T}_+ \odot \mathbf{R}_+^H) \right)^{-1}_K \right) \\ &\succeq \frac{\sigma^2}{2T} \left(\Re(\mathbf{T} \odot \mathbf{R}_S^H) \right)^{-1} = \text{CRB}_{Ny}. \end{aligned} \quad (38)$$

(38) shows that the estimate performance for $\phi = [\phi_1, \dots, \phi_K]$ in the scene where there are only K signals from ϕ is better than that in the scene where there are both the K signals from ϕ and the increased signal from ϕ_{K+1} . In other words, the increase of the number of DOA will degrade the performance of DOA estimate. It is simplistic to conclude that the estimation variance is lowest when there is only one signal. After calculation, the lowest estimation variance is

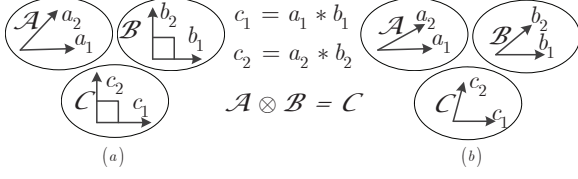


Fig. 2. 2D union space.

$\frac{6}{\text{SNR} \cdot T \cdot M(M^2-1)}$, where signal-to-noise ratio (SNR) is defined as $\text{SNR} = (E(|s(t)|^2)/\sigma^2)$.

However, the performance of DOA estimation based on the proposed model will not degrade with the increase of the number of DOAs. It's easy to get $\mathbf{E} = \mathbf{D} * \mathbf{B}_\Omega$ since $\mathbf{G}_S = \mathbf{A} * \mathbf{B}_\Omega$. We further hold that

$$\mathbf{G}_S^H \mathbf{G}_S = (\mathbf{A}^H \mathbf{A}) \odot (\mathbf{B}_\Omega^H \mathbf{B}_\Omega), \quad (39)$$

$$\mathbf{E}^H \mathbf{E} = (\mathbf{D}^H \mathbf{D}) \odot (\mathbf{B}_\Omega^H \mathbf{B}_\Omega). \quad (40)$$

Thus, $\mathbf{G}_S^H \mathbf{G}_S = (\mathbf{A}^H \mathbf{A}) \odot \mathbf{I}$, $\mathbf{E}^H \mathbf{E} = (\mathbf{A}^H \mathbf{A}) \odot \mathbf{I}$ hold when $P = L$. Based on above results and after some matrix manipulations, we know that the estimation variance based on our model maintains at $\frac{6}{\text{SNR} \cdot T \cdot M(M^2-1)}$ all the time when $P = L$. Consequently, when $L = P$, we have

$$\begin{aligned} \text{CRB}_{\text{sub}} &= \text{CRB}_{\text{Ny}}, \text{ when } K = 1, \\ \text{CRB}_{\text{sub}} &\leq \text{CRB}_{\text{Ny}}, \text{ when } K > 1. \end{aligned} \quad (41)$$

Remark 2. We can view (41) from the physical perspective. If we study the cross-correlation δ_{ij} between \mathbf{B}^i and \mathbf{B}^j , i.e. $\delta_{ij} = |(\mathbf{B}^i)^H \mathbf{B}^j|$, $i \neq j$, it will be easy to obtain

$$\begin{cases} \delta_{ij} < 1, P < L \\ \delta_{ij} = 0, P = L \end{cases}. \quad (42)$$

Since (39) and (42) hold, the cross-correlation of the new steer vectors \mathbf{G}_S is lower than that of the primary steer vectors \mathbf{A} no matter whether P is equal to L or not. Specifically, the new steer vectors are completely uncorrelated when $P = L$ in spite of the primary steer vectors are correlated. At this time, the new DOA estimation is equivalent to execution one by one in a scene where there is only one signal. This is an explanation why the performance of DOA estimation based on our model will not degrade with the increasing of the DOA number while the performance of DOA estimation based on the primary model will degrade. From the perspective geometry, this situation is corresponding to Fig.2 (a), where as long as the vectors in space \mathcal{B} are orthogonal, vectors in space \mathcal{C} will be orthogonal no matter whether the vectors in space \mathcal{A} are orthogonal, even if they are same. A more general case i.e. $P < L$, is corresponding to Fig.2 (b), where even if the vectors in space neither \mathcal{A} nor \mathcal{B} are orthogonal, vectors in space \mathcal{C} will be less correlated, i.e., the angle between the vectors in space \mathcal{C} is larger than before. In an subjective sense, the cross-correlation of steer vectors reflects similarity and identifiability of the DOA. The lower cross-correlation of steer vectors, the easier they are to be distinguished.

VI. SIMULATION

In this section, we present some numerical simulations to illustrate the performance of the proposed algorithms. In our examples, we consider some complex-valued narrowband far-field non-coherent signals with equal power impinging on a ULA composed of $M = 8$ sensors which are separated by a half wavelength corresponding to Nyquist sampling rate, which would probably be the signal highest frequency. In the simulations, we fix the number of snapshots at $T = 4000$ for Nyquist sampling, $T_{\text{sub}} = T/L$ for sub-Nyquist sampling, the Nyquist sampling rate at $f_N = 10$ GHz, and the sampling rate reduction factor at $L = 20$.

The root-mean-square error (RMSE) of DOA is defined as $\text{RMSE} = \sqrt{\frac{1}{N_m K} \sum_{i=1}^{N_m} \sum_{k=1}^K (\theta_k^i - \hat{\theta}_k^i)^2}$, where the superscript i refers to the i th trial, N_m denotes the number of Monte Carlo tests. And the definitions of RMSE of spatial phase and frequency are similar to that of DOA.

Later on, we will study the performance versus different noise levels, different branch number, or different source number. We will compare our methods with ST-Euler-ESPRIT in [18]. The receiver configuration parameters of ST-Euler-ESPRIT are the same with ours. The delay is Nyquist sampling interval $T_N = 1/f_N$. Hereon, we give the reasons why we choose ST-Euler-ESPRIT: i) In [15], both frequency and DOA estimation accuracy are low since they are limited by the reciprocal of block length and the array aperture, respectively. ii) In terms of the hardware complexity, [17] and [16] is the simplified version [18], and [18] has the same hardware complexity with our receiver. For the sake of fair, we compare our methods with [18]. 20000 Monte Carlo trials for each example are implemented in this section.

A. Performance with noise

In this subsection, it will be shown that our model can be solved by the proposed algorithm in different noise levels. We add an array construction MRA to prove the validity of the algorithm in this subsection. The MRA is composed of $M = 8$ sensors which are located at $\mathbf{d} = [0, 1, 4, 10, 16, 22, 28, 30]d$. However ST-Euler-ESPRIT will be feasible only when ULA is employed.

In this subsection, we set the branch number $P = L$. Meanwhile, the average sampling rate $f_A = \frac{P f_N}{L}$ is equal to the Nyquist sampling rate. We consider that the signal number $K = 3$, and signals are from $\boldsymbol{\theta} = [\theta_1, \theta_2, \theta_3]$, where $\theta_1, \theta_2, \theta_3$ are subject to $[-12.5^\circ, -7.5^\circ]$, $[-2.5^\circ, 2.5^\circ]$, $[7.5^\circ, 12.5^\circ]$ uniformly distribution, respectively. And the signal carrier frequency $\mathbf{f} = [f_1, f_2, f_3]$ are subject to $[0.5, 9.5]$ GHz uniformly distribution, and any two signals are not in the same sub-band.

Fig.3-Fig.5 depict the RMSE versus SNR in terms of spatial phase, frequency, and DOA estimation, respectively. Fig.3 shows that the phase estimation performance of algorithms JDFTD and JDFSD improves with SNR and achieves the CRB_{sub} when ULA or MRA is employed. Although ST-Euler-ESPRIT has a similar trend, the performance is inferior to JDFTD and JDFSD when ULA is employed. And we also find that the phase estimation performance of JDFTD and JDFSD will improve obviously when MRA is employed since the

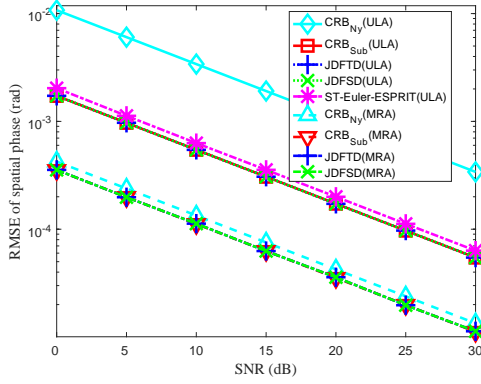


Fig. 3. RMSE of phase estimates versus SNR.

MRA widens the array aperture. But ST-Euler-ESPRIT can not employ the MRA as explained previously. As expected, Fig.3 shows that CRB_{sub} is lower than CRB_{Ny} when $L = P$ and $K > 1$. This illustrates that our algorithms for the new model can obtain a better phase estimation than the Nyquist sampling structure.

Fig.4 demonstrates that the frequency estimation performance of JDFTD and JDFSD can achieve the CRB_{sub} . But the frequency estimation performance of ST-Euler-ESPRIT is obviously worse than JDFTD and JDFSD. From the view of frequency estimation, employing the array receiving is equal to enhancing the SNR. So the frequency estimation performance is the same when the sensor number is the same whether the array is ULA or MRA. Combined with the conclusion in [25], the CRB_{sub} and CRB_{Ny} for frequency estimation are given as follow when T is sufficiently large.

$$\text{CRB}_{\text{sub}}(f) = \frac{1}{4\pi^2} \frac{6}{\text{SNR}} \frac{1}{M} \frac{f_N^2}{T^3} \frac{L}{P}. \quad (43)$$

$$\text{CRB}_{\text{Ny}}(f) = \frac{1}{4\pi^2} \frac{6}{\text{SNR}} \frac{1}{M} \frac{f_N^2}{T^3}. \quad (44)$$

It is particularly obvious that $\text{CRB}_{\text{sub}}(f) = \text{CRB}_{\text{Ny}}(f)$ when $L = P$ according to (43).

Comparing Fig.5 with Fig.3, it is concluded that the performances of DOA estimation and phase estimation have the same trend. Because the sampling number M in space domain is much less than the sampling number T in time domain, the phase estimation is worse than the frequency estimation. Based on (2), we know that the performance of DOA estimation is mainly influenced by the phase estimation. Because of this, we will only give the phase estimation simulation result rather than the DOA estimation simulation result in the following simulations.

B. Performance with various branch number

In this subsection, we will investigate the estimation performance in the case of different branch number. The simulation conditions are the same with that in subsection VI-A except that the branch number P changes from 4 to 20 at 2 interval - that is, the average sampling rate $f_A = \frac{Pf_N}{L}$ changes from $0.2f_N$ to f_N at $0.1f_N$ interval. In this simulation, we employ a random sampling pattern C .

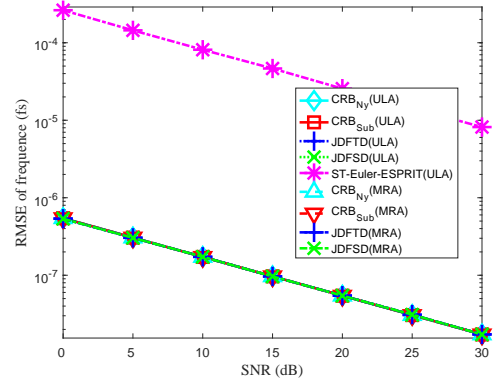


Fig. 4. RMSE of frequency estimates versus SNR.

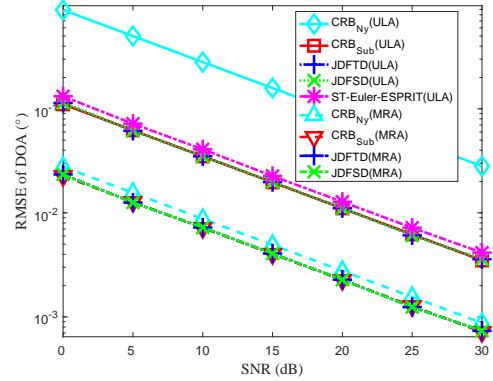


Fig. 5. RMSE of DOA estimates versus SNR.

Fig.6 shows that the phase (DOA) estimation performances of algorithm JDFTD and JDFSD improve with the branch number and reach the CRB_{sub} when P is relatively large. JDFTD is slightly worse than JDFSD or CRB_{sub} . This is because that the cross-correlation of the column vectors of \mathbf{B} is obviously great when P is remarkably small and this leads to trilinear decomposition slow convergence. It is not surprising that the CRB_{sub} is lower than CRB_{Ny} when $L = P$. But we notice that this phenomenon still exists even $P = 0.2L$. This illustrates that the benefit from the decrease of the cross-correlation of steer vectors from (42) is much larger than the loss caused by the decrease of samplings. Namely, we can realize a much better phase and DOA estimation performance with much less samplings. Fig.6 also shows that both JDFTD and JDFSD are superior to ST-Euler-ESPRIT with any branch numbers, especially with small numbers.

Fig.7 shows that the frequency estimation performances of algorithm JDFTD and JDFSD are improved with the branch number and achieve the CRB_{sub} when P is relatively large. The performances of JDFTD and JDFSD are slightly worse than the CRB_{sub} . Obviously, CRB_{sub} is higher than CRB_{Ny} except $P = L$. However, the frequency estimation performance of ST-Euler-ESPRIT is obviously worse than JDFTD and JDFSD no matter how many branches there are.

We notice that when the average sampling rate is lower than the Nyquist sampling rate where $P < L$, the CRB for our model is lower than the conventional CRB in terms of

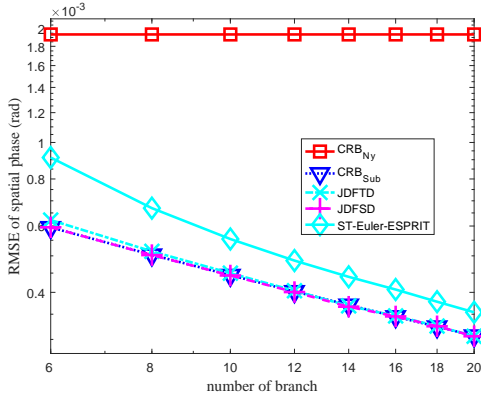


Fig. 6. RMSE of phase estimates versus number of branch.

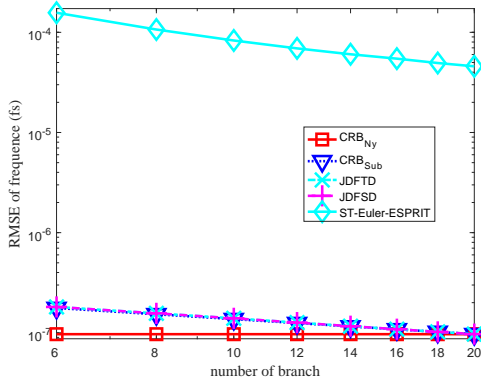


Fig. 7. RMSE of frequency estimates versus number of branch.

DOA or spatial phase estimation, and the opposite happens in terms of frequency estimation. However, the performance of DOA or spatial phase estimation is usually far worse than that of frequency estimation because of $M \ll T$. So, we care more about the performance improvement of DOA or spatial phase estimation than that of frequency estimation. Namely, comparing with the performance improvement of DOA or spatial phase estimation, the performance degradation of frequency estimation is insignificant.

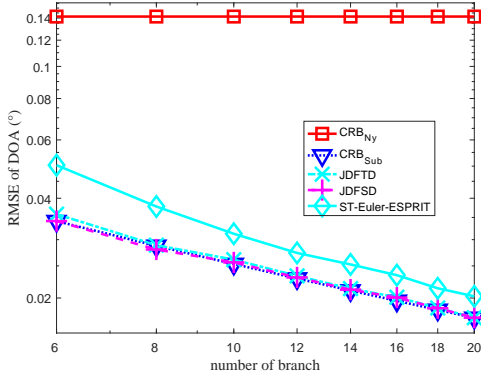


Fig. 8. RMSE of DOA estimates versus number of branch.

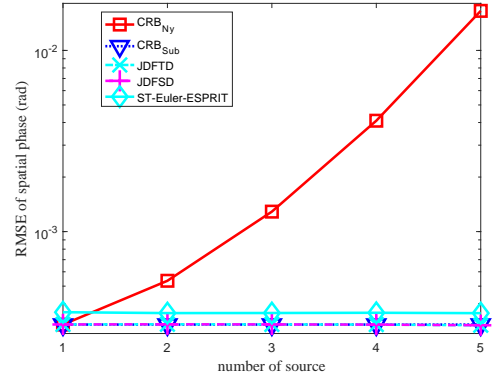


Fig. 9. RMSE of phase estimates versus number of source.

C. Performance with various signal number

In this subsection, we will investigate the estimation performance when the signal number changes. The simulation conditions are the same with subsection VI-A except that the signal number K changes from 1 to 5 one by one and only ULA is considered. We set $\vartheta = [\vartheta_1, \vartheta_2, \dots, \vartheta_5]$, where ϑ_i is subject to $[10j - 2.5^\circ, 10j + 2.5^\circ]$ uniformly distribution, where $i = -2, -1, \dots, 2$, $j = -2, -1, \dots, 2$ and i has no corresponding relationship with j . We set DOA $\theta = [\vartheta_1, \vartheta_2, \dots, \vartheta_K]$. Let $\mathbf{v} = [v_1, v_2, \dots, v_5]$ are subject to $[0.5, 9.5]$ GHz uniformly distribution, at the same time any two signals are not in the same one sub-band. The signal carrier frequency is set at $\mathbf{f} = [v_1, v_2, \dots, v_K]$.

Fig.9 shows that the phase (DOA) estimation performances of algorithm JDFTD and JDFSD are not influenced by the signal number and maintain CRB_{sub} . However, CRB_{Ny} increases with the signal number, and increases faster than exponential function of the signal number. Besides, CRB_{Ny} is equal to CRB_{sub} only when $K = 1$, otherwise CRB_{Ny} is higher than CRB_{sub} . These results meet the analysis in section V. As for ST-Euler-ESPRIT, although it is also negligibly affected by the signal number, its phase (DOA) estimation performance is worse than JDFTD and JDFSD.

Fig.10 shows that the frequency estimation performances of algorithm JDFTD and JDFSD are not influenced by the signal number and can reach CRB_{sub} . At the same time, CRB_{Ny} is equal to CRB_{sub} because $P = L$. The frequency estimation performance of ST-Euler-ESPRIT is particularly worse than that of JDFTD and JDFSD as before.

VII. CONCLUSIONS

In this paper, for the scenario where several narrowband far-field signals whose carrier frequencies are far separated impinging on an array, we designed an array receiver architecture by introducing the sub-Nyquist sampling technology. We derived a time-space union signal reception model with taking the spatial sampling and sub-Nyquist sampling into consideration simultaneously. Meanwhile, we can decrease the time-domain sampling rate and improve the DOA estimation accuracy.

We proposed two joint DOA and frequency estimation algorithms for this model, one is based on trilinear decomposition

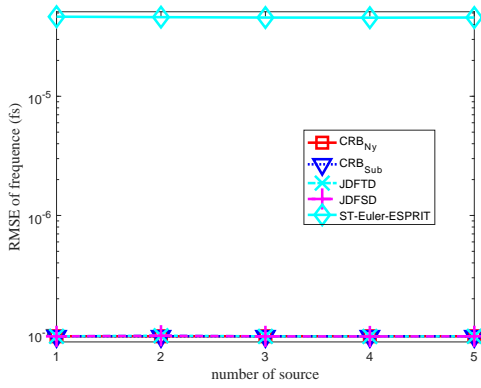


Fig. 10. RMSE of frequency estimates versus number of source.

from the perspectives of third-order tensor and the other is based on subspace decomposition.

In terms of spatial phase estimation, we derived the CRB for the model, and proved that the CRB is immune to the signal number when the branch number of our architecture is equal to the sampling rate reduction factor, and is lower than the CRB for the conventional model which employs Nyquist sampling. Furthermore, the new steer vectors are completely uncorrelated under the limited number of sensors, which makes a big improvement for the spatial phase estimation performance. From the geometry perspective, the estimation performance improvement benefits from the union time-space model.

The simulations validated that the receiver architecture and the proposed approaches are feasible, and the variances of the proposed approaches are very close to their CRB and are beyond the CRB which employs Nyquist sampling in the case of different noise levels, different branch number, or different source number. Specifically, the variances of the proposed approaches are lower than the CRB which employs Nyquist sampling when the branch number of our architecture is less than the sampling rate reduction factor—that is, the average sampling rate is lower than the Nyquist sampling rate.

REFERENCES

- [1] S. Haykin, "Cognitive radio: brain-empowered wireless communications," *IEEE J. Sel. Areas Commun.*, vol. 23, no. 2, pp. 201–220, Feb 2005.
- [2] T. Yucek and H. Arslan, "A survey of spectrum sensing algorithms for cognitive radio applications," *IEEE Commun. Surveys Tuts.*, vol. 11, no. 1, pp. 116–130, First 2009.
- [3] M. Mishali and Y. C. Eldar, "Wideband spectrum sensing at sub-Nyquist rates [applications corner]," *IEEE Signal Process. Mag.*, vol. 28, no. 4, pp. 102–135, July 2011.
- [4] H. Sun, A. Nallanathan, C. X. Wang, and Y. Chen, "Wideband spectrum sensing for cognitive radio networks: a survey," *IEEE Wireless Commun.*, vol. 20, no. 2, pp. 74–81, April 2013.
- [5] D. Cohen and Y. C. Eldar, "Sub-Nyquist sampling for power spectrum sensing in cognitive radios: A unified approach," *IEEE Trans. Signal Process.*, vol. 62, no. 15, pp. 3897–3910, Aug 2014.
- [6] H. Krim and M. Viberg, "Two decades of array signal processing research: the parametric approach," *IEEE Signal Process. Mag.*, vol. 13, no. 4, pp. 67–94, Jul 1996.
- [7] R. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE Trans. Antennas Propag.*, vol. 34, no. 3, pp. 276–280, Mar 1986.
- [8] R. Roy, A. Paulraj, and T. Kailath, "ESPRIT—a subspace rotation approach to estimation of parameters of cisoids in noise," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 34, no. 5, pp. 1340–1342, Oct 1986.
- [9] A. N. Lemma, A. J. van der Veen, and E. F. Deprettere, "Joint angle-frequency estimation using multi-resolution esprit," in *Proc. IEEE Int. Conf. on Acous. Speech and Signal Process. (ICASSP)*, vol. 4, May 1998, pp. 1957–1960 vol.4.
- [10] A. Lemma, A. J. V. der Veen, and E. Deprettere, "Analysis of joint angle-frequency estimation using esprit," *IEEE Trans. Signal Process.*, vol. 51, no. 5, pp. 1264–1283, May 2003.
- [11] M. Mishali and Y. C. Eldar, "Sub-Nyquist sampling," *IEEE Signal Process. Mag.*, vol. 28, no. 6, pp. 98–124, Nov 2011.
- [12] —, "From theory to practice: Sub-Nyquist sampling of sparse wide-band analog signals," *IEEE J. Sel. Topics Signal Process.*, vol. 4, no. 2, pp. 375–391, April 2010.
- [13] Y. C. Eldar and T. Michaeli, "Beyond bandlimited sampling," *IEEE Signal Process. Mag.*, vol. 26, no. 3, pp. 48–68, May 2009.
- [14] M. Mishali and Y. C. Eldar, "Blind multiband signal reconstruction: Compressed sensing for analog signals," *IEEE Trans. Signal Process.*, vol. 57, no. 3, pp. 993–1009, March 2009.
- [15] D. D. Ariananda and G. Leus, "Compressive joint angular-frequency power spectrum estimation," in *Proc. Eur. Sig. Process. Conf. (EUSIPCO)*, Sep 2013, pp. 1–5.
- [16] A. A. Kumar, S. G. Razul, and C. M. S. See, "An efficient sub-Nyquist receiver architecture for spectrum blind reconstruction and direction of arrival estimation," in *Proc. IEEE Int. Conf. on Acous. Speech and Signal Process. (ICASSP)*, May 2014, pp. 6781–6785.
- [17] —, "Spectrum blind reconstruction and direction of arrival estimation at sub-Nyquist sampling rates with uniform linear array," in *Proc. IEEE Int. Conf. Digital Sig. Process. (DSP)*, July 2015, pp. 670–674.
- [18] A. A. Kumar, S. G. Razul, M. G. Chandra, C. M. See, and P. Balamuralidhar, "Joint frequency and direction of arrival estimation with space-time array," in *IEEE Sens. Array and Multichannel Signal Process. Workshop (SAM)*, July 2016, pp. 1–5.
- [19] S. Stein, O. Yair, D. Cohen, and Y. C. Eldar, "Joint spectrum sensing and direction of arrival recovery from sub-Nyquist samples," in *Proc. IEEE Signal Process. Adv. Wireless Commun. (SPAWC)*, June 2015, pp. 331–335.
- [20] J. B. Kruskal, "Three-way arrays: rank and uniqueness of trilinear decompositions, with application to arithmetic complexity and statistics," *Linear Algebra Appl.*, vol. 18, no. 2, pp. 95 – 138, 1977.
- [21] L. D. Lathauwer, "A link between the canonical decomposition in multilinear algebra and simultaneous matrix diagonalization," *SIAM J. Matrix Anal. A.*, vol. 28, no. 3, pp. 642–666, 2006.
- [22] L. De Lathauwer and D. Nion, "Decompositions of a higher-order tensor in block terms-part III: Alternating least squares algorithms," *SIAM J. Matrix Anal. A.*, vol. 30, no. 3, pp. 1067–1083, 2008.
- [23] N. Li, S. Kindermann, and C. Navasca, "Some convergence results on the regularized alternating least-squares method for tensor decomposition," *Linear Algebra Appl.*, vol. 438, no. 2, pp. 796 – 812, 2013.
- [24] C. Navasca, L. D. Lathauwer, and S. Kindermann, "Swamp reducing technique for tensor decomposition," in *Proc. Eur. Sig. Process. Conf. (EUSIPCO)*, Aug 2008, pp. 1–5.
- [25] D. C. Rife and R. R. Boorstyn, "Single tone parameter estimation from discrete-time observations," *IEEE Trans. Inf. Theory*, vol. 20, no. 5, pp. 591–598, 1974.
- [26] P. Stoica and A. Nehorai, "MUSIC, maximum likelihood, and cramer-rao bound: further results and comparisons," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 38, no. 12, pp. 2140–2150, 1990.
- [27] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge University Press, 1985.